



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

TET1 interacts directly with NANOG via independent domains containing hydrophobic and aromatic residues

Citation for published version:

Pantier, R, Mullin, NP, Hall-Ponsele, E & Chambers, I 2020, 'TET1 interacts directly with NANOG via independent domains containing hydrophobic and aromatic residues', *Journal of Molecular Biology*, vol. 432, no. 23, pp. 6075-6091. <https://doi.org/10.1016/j.jmb.2020.10.008>

Digital Object Identifier (DOI):

[10.1016/j.jmb.2020.10.008](https://doi.org/10.1016/j.jmb.2020.10.008)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Journal of Molecular Biology

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





TET1 Interacts Directly with NANOG via Independent Domains Containing Hydrophobic and Aromatic Residues

Raphaël Pantier, Nicholas Mullin, Elisa Hall-Ponsele and Ian Chambers*

Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, Scotland, United Kingdom

Correspondence to Ian Chambers: ichambers@ed.ac.uk (I. Chambers)

<https://doi.org/10.1016/j.jmb.2020.10.008>

Edited by Sepideh Khorasanizadeh

Abstract

The DNA demethylase TET1 is highly expressed in embryonic stem cells and is important both for lineage commitment, and reprogramming to naive pluripotency. TET1 interacts with the pluripotency transcription factor NANOG which may contribute to its biological activity in pluripotent cells. However, how TET1 interacts with other proteins is largely unknown. Here, we characterise the physical interaction between TET1 and NANOG using embryonic stem cells and bacterial expression systems. TET1 and NANOG interact through multiple binding sites that act independently. Critically, mutating conserved hydrophobic and aromatic residues within TET1 and NANOG abolishes the interaction. On chromatin, NANOG is predominantly localised at ESC enhancers. While TET1 binds to CpG dinucleotides in promoters using its CXXC domain, TET1 also binds to enhancers, though the mechanism involved is unknown. Comparative ChIP-seq analysis identifies genomic loci bound by both TET1 and NANOG, that correspond predominantly to pluripotency enhancers. Importantly, around half of NANOG transcriptional target genes are associated with TET1-NANOG co-bound sites. These results indicate a mechanism by which TET1 protein may be targeted to specific sites of action at enhancers by direct interaction with a transcription factor.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Ten-eleven-translocation (TET) family proteins are responsible for active DNA demethylation by oxidation of 5-methylcytosine^{1,2} and play important roles during embryonic development and various physiological processes.³ TET proteins contribute to DNA demethylation in naive embryonic stem cells (ESCs),^{4–7} in particular at enhancers.^{8–13} TET protein activity is required both for proper differentiation^{14,15} and reprogramming to pluripotency.^{16–18} TET1 is the most highly expressed TET family protein both in pluripotent cells and during early development.^{19–21} TET1 predominantly binds to promoters via its N-terminal CXXC domain which recognises unmethylated CpG dinucleotides.^{22–25} TET1 binding at enhancers in

ESCs^{26–28} could be mediated by interactions with the pluripotency factors NANOG, PRDM14, OCT4 and SOX2.^{29–32} Interestingly, co-expression of TET1 and NANOG in pre-iPS cells synergistically enhances reprogramming to pluripotency.²⁹ However, how TET1 might be recruited to chromatin via protein–protein interactions remains poorly understood with little known about the residues involved in protein binding.

Here, the interaction between TET1 and the pluripotency factor NANOG was characterised in ESCs. Co-immunoprecipitations using an array of TET1 truncations and mutants uncovered novel regions involved in protein–protein interactions, both within and outwith the well characterised catalytic domain. Furthermore, alanine mutagenesis identified single residues that show

high evolutionary conservation and that contribute to the interaction of TET1 with NANOG. Comparison of TET1 and NANOG ChIP-seq datasets identified genomic loci that are putatively regulated by the TET1-NANOG complex.

Results

The TET1 N-terminus interacts directly with NANOG via the evolutionary conserved residues L110 and L114

The TET1 protein expressed in mouse ESCs is composed of 2039 residues. TET1 is characterised by an evolutionary conserved C-terminal catalytic domain, that can be subdivided into a cysteine rich region (residues 1367–1550) and a double stranded beta helix domain (DSBH) (residues 1551–2039) (Figure S1(a)). TET1 also possesses a CXXC domain (residues 567–608), a DNA binding region.³³ NANOG is a 305 amino acids transcription factor comprising a N-terminal domain (residues 1–95), a DNA binding homeodomain (residues 96–155) and a C-terminal region containing a tryptophan-repeat (WR) (residues 199–243) (Figure S1(a)). TET1 has been identified as a NANOG-binding protein by independent affinity purification-mass spectrometry analyses.^{29,30} We therefore analysed the interaction between endogenous TET1 and NANOG in pluripotent cells using nuclear protein extracts from *Tet1*-(Flag)₃ ESCs²¹ immunoprecipitated with an anti-Flag antibody. Relative to controls from E14Tg2a ESCs which showed only background binding, FLAG immunoprecipitates from *Tet1*-(Flag)₃ ESCs were strongly enriched for NANOG (Figure S1(b)). This confirms that NANOG and TET1 proteins interact in ESCs when expressed at endogenous levels. Notably, TET1 protein in immunoprecipitates migrates slower than input material on immunoblots; a phenomenon previously reported.^{34–36} Next, to determine whether TET1 interacts with NANOG via the TET1 N- or C-terminus, two large (Flag)₃-tagged TET1 fragments 1–631 and 734–2039 were cloned and expressed in ESCs (Figure 1(a)), together with NANOG. Following TET1 immunoprecipitation, NANOG was co-immunoprecipitated with both constructs (Figure 1(b)). As TET1 1–631 and 734–2039 do not contain overlapping residues, these results suggest that TET1 contains at least two NANOG-interacting domains (NIDs) that function independently.

To begin to explore how NANOG binds to the TET1 N-terminus, (Flag)₃-tagged TET1 fragments 1–321, 1–215 and 1–108 were cloned and expressed in ESCs (Figure S1(c)), together with NANOG. NANOG was co-immunoprecipitated with TET1(1–321) and TET1(1–215) but not TET1(1–108) (Figure S1(d)). To home in on the NANOG-interacting domain within the TET1 N-terminus and to determine whether the

interaction between the TET1 N-terminus and NANOG was direct, both (His)₆-tagged NANOG and several MBP-tagged TET1 fragments (1–321, 1–215, 1–165, 1–120, 1–108) were cloned into IPTG-inducible plasmids and expressed in *Escherichia coli* (Figure 1(c)). MBP-TET1 fragments purified using an amylose resin were examined for co-purifying NANOG by immunoblotting. NANOG co-purified with all TET1 fragments, except TET1(1–108) which showed a dramatically decreased interaction with NANOG (Figure 1(d)). Importantly, these experiments confirmed a direct physical interaction and narrowed down the first NANOG-interacting domain (NID 1) to 11 residues (109–120). Protein alignments showed that NID 1 is highly conserved among mammals, indicating a selective pressure for the conservation of these TET1 residues (Figure S1(e)). However, residues 109–120 of TET1 do not align with TET2 or TET3 proteins (data not shown). To identify which residues are responsible for binding to NANOG, the MBP-TET1 1–120 plasmid construct was modified by alanine substitution of specific amino acids (proline, glutamine, arginine, leucine, serine and valine) within residues 109–120 (Figure 1(e)). The binding of NANOG to each mutant construct was assessed following bacterial expression and TET1 purification. Strikingly, only the L→A mutant (L110A, L114A) showed a decreased interaction with NANOG, which was reduced to a similar extent as the negative control TET1 1–108 (Figure 1(f)). Together, these data indicate that one or both of the two evolutionary conserved leucine residues (L110 and/or L114) are in direct physical interaction with NANOG.

To determine which regions of NANOG interact directly with the TET1 N-terminus, MBP-TET1 1–321 was co-expressed with several (His)₆-tagged NANOG truncations in *E. coli* (Figure 2(a)). Surprisingly, the TET1 N-terminus interacted with three out of four NANOG truncations: 1–160, 91–246, and 194–305 (Figure 2(b)). However, the TET1 N-terminus showed no interaction with the NANOG homeodomain (Figure 2(b)). Moreover, while the NANOG WR interacts with MBP-SOX2 (positive control³⁰), MBP-TET1 1–321 showed no physical interaction with the NANOG WR (Figure S2). Collectively, these results indicate that the TET1 N-terminus interacts with several independent sites on NANOG, and independently of its two most characterised domains: the homeodomain and the WR (Figure 2(c)).

The TET1 C-terminus contains two domains that bind NANOG via aromatic interactions

Following the initial observation that TET1 contains >1 independent NANOG-interacting domains (Figure 1(b)), the TET1 C-terminus was analysed to identify the NANOG-interacting residues. Regions of TET1 extending from 734 to

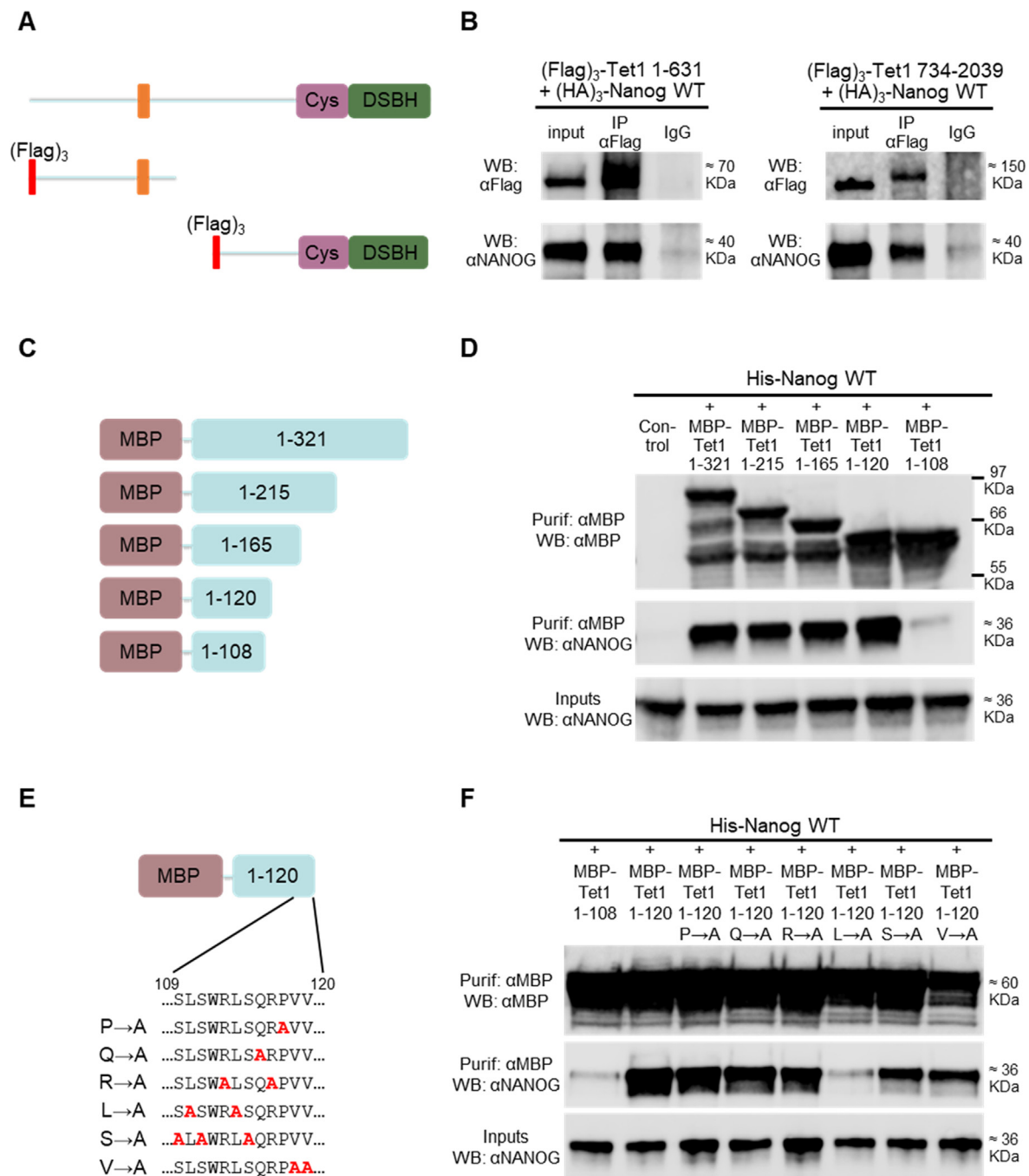


Figure 1. TET1 contains independent NANOG-binding regions. (a) and (b) Co-immunoprecipitations of non-overlapping (Flag)₃-TET1 N- and C-terminal constructs (a) with (HA)₃-NANOG from E14/T ESCs. (b) Immunoblots were probed with the antibodies indicated on the left (representative images, *n* = 2). (c) and (d) Co-purification of MBP-tagged TET1 N-terminal constructs with (His)₆-NANOG from *E. coli*. (c) Fragments of the TET1 N-terminus are shown in the context of full length TET1. (d) Immunoblots were probed with the antibodies indicated on the left (representative images, *n* = 3). (e) and (f) Co-purification of (His)₆-NANOG with alanine substitution mutants of MBP-TET1 (1–120) from *E. coli*. (e) Alanine substitution mutants of TET1 (109–120) are shown in the context of full length TET1. (f) Immunoblots were probed with the antibodies indicated on the left (representative images, *n* = 3).

varying degrees towards the Cys domain were expressed together with NANOG in ESCs (Figures 3(a) and S3(a)). TET1 fragments containing truncations up to residue 1181 (734–

1229, 734–1202, 734–1181) were able to bind NANOG, while further C-terminal truncations (734–1155 and 734–1131) abolished interaction with NANOG (Figure S3(a) and (b)). Analysis of a

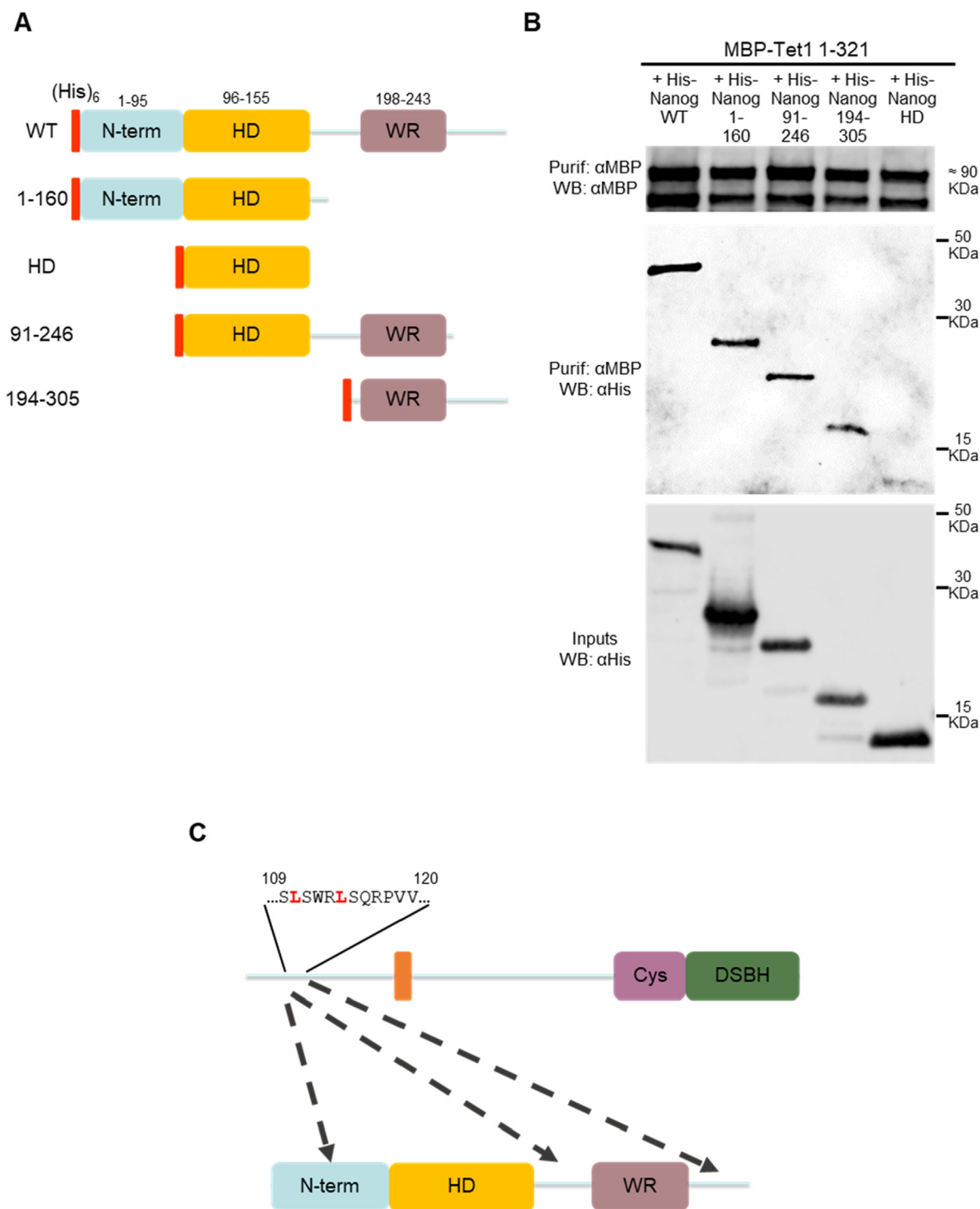


Figure 2. Multiple NANOG regions interact with the TET1 N-terminal domain. (a) and (b) Co-purification of MBP-TET1 1–321 with (His)₆-NANOG constructs (a) from *E. coli*. (b) Immunoblots were probed with the antibodies indicated on the left (representative images, *n* = 2). (c) Diagram of the interactions between the TET1 NID 1 and NANOG, highlighting critical leucines (red). Dashed arrows indicate potential TET1-interacting regions in NANOG.

further truncation (734–1169) narrowed down the second NANOG-interacting domain of TET1 to residues 1156–1169 (NID 2) (Figure 3(a) and (b)). A construct containing NID 2 did not interact with NANOG when co-expressed in *E. coli* (Figure S3

(c) and (d)) suggesting that either the interaction observed in ESCs is indirect, or that a direct interaction dependent on post-translational modifications and/or protein folding could not be reproduced in bacteria. However, this interaction

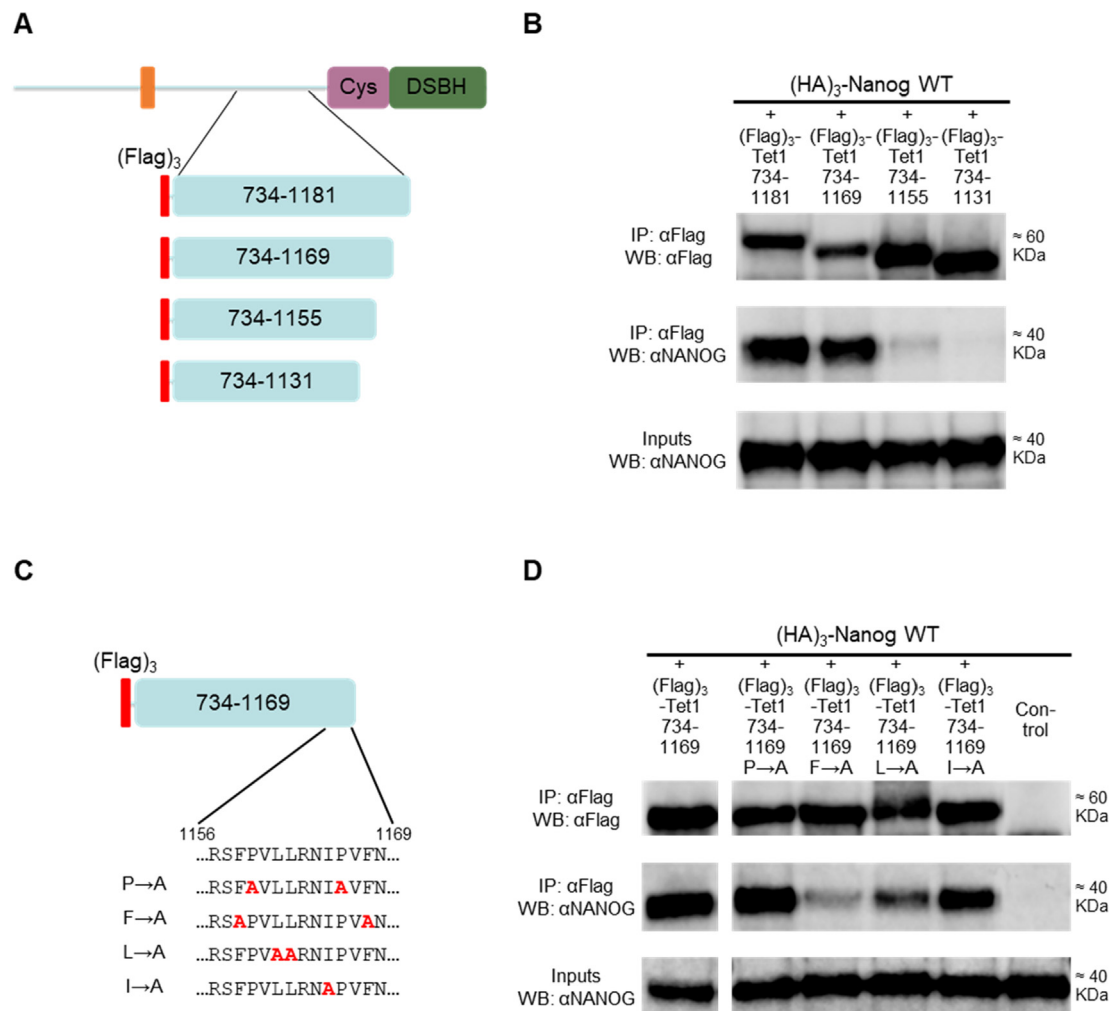


Figure 3. The NANOG interaction domain 2 of TET1 binds NANOG using phenylalanine residues. (a) and (b) Co-immunoprecipitation of the (Flag)₃-TET1 constructs (a) with (HA)₃-NANOG from E14/T ESCs. (b) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 3$). (c) and (d) Co-immunoprecipitation of alanine substitution mutants of (Flag)₃-TET1(734–1169) with (HA)₃-NANOG from E14/T ESCs. (c) Alanine substitutions of the amino acids within residues 1156–1169. (d) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 2$).

does not depend on phosphorylation. While treatment of ESC protein extracts with phosphatase affected the mobilities of both TET1 and NANOG proteins, TET1(734–1169) retained the capacity to bind NANOG (Figure S3(e)).

Residues 1156–1169 of mouse TET1 (NID 2) have a high similarity to sequences of TET1 proteins from other mammals, with phenylalanine 1158 strictly conserved (Figure S3(f)). This region did not align with TET2 or TET3 proteins (data not shown). To identify residues that bind NANOG, the expression plasmid encoding TET1 734–1169 was modified by alanine substitution of specific amino acids (proline, phenylalanine, leucine and isoleucine) within residues 1156–1169 (Figure 3 (c)). The F→A mutant (F1158A, F1168A) was the only mutant that showed a decreased interaction

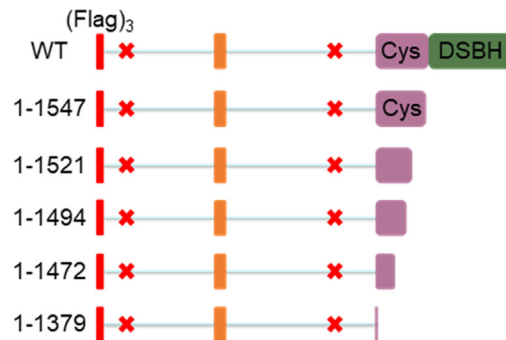
with NANOG (Figure 3(d)). Together, these results indicate that phenylalanine 1158 and/or 1168 are critical for NANOG binding.

The preceding results identified two independent NANOG-interacting domains within residues 109–120 (NID 1) and 1156–1169 (NID 2), respectively in the N- and C-terminal fragments. TET1 fragments containing deletions of these regions did not interact with NANOG compared to the unmutated version (Figure S4(a) and (b)). Full-length TET1 constructs with deletions in NID 1 ($\Delta 1$), NID 2 ($\Delta 2$) or both ($\Delta 1+2$) were therefore generated (Figure S4(c)). A TET1 mutant lacking a low-complexity insert ($\Delta 1733-1901$) was used as a control (Figure S4(c)), as this region has been hypothesised to function in protein–protein interactions.³⁷ As expected, NANOG was co-

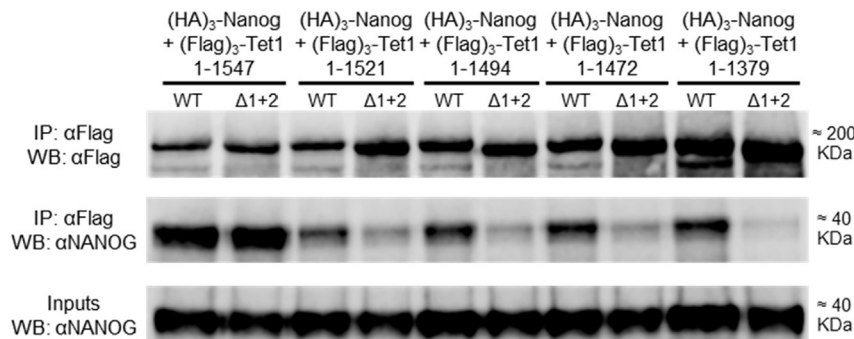
immunoprecipitated with each of the single TET1 mutants. Surprisingly however, although the TET1 $\Delta 1+2$ double-mutant showed reduced NANOG binding compared to wild-type, binding was not completely eliminated (Figure S4(d)). This suggests that an additional NANOG-interacting domain may

exist in TET1. To identify this third NID, TET1 expression plasmids were generated combining the double mutation $\Delta 1+2$ with increasing C-terminal truncations (Figure 4(a)). Plasmids with wild-type TET1 coding sequence used as controls allowed assessment of the relative importance of

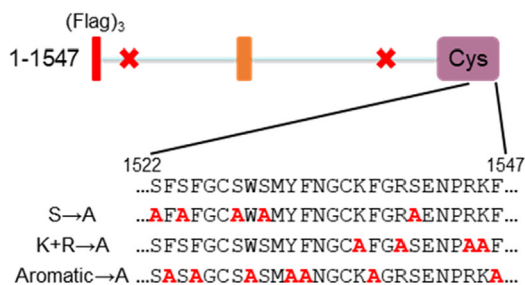
A



B



C



D

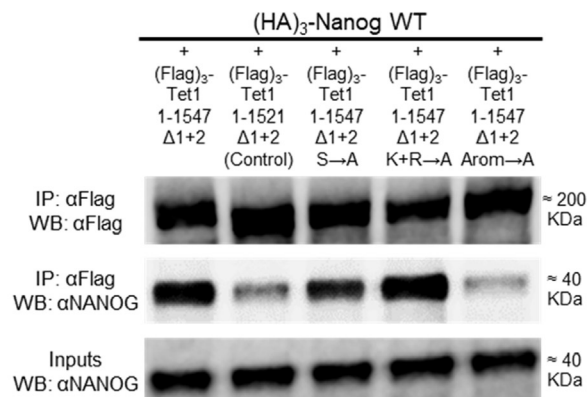


Figure 4. A third NANOG-interacting domain overlapping TET1 catalytic domain. (a) and (b) Co-immunoprecipitation of (Flag)₃-TET1 truncations with (HA)₃-NANOG in E14/T ESCs. (a) TET1 truncations were prepared in parallel in an unmutated (WT) Tet1 plasmid (not shown for simplicity) or one carrying the $\Delta 109-120 + \Delta 1132-1202$ ($\Delta 1+2$) mutations (red crosses). (b) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 3$). (c) and (d) Co-immunoprecipitations of (Flag)₃-TET1 mutants with (HA)₃-NANOG from E14/T ESCs. (c) Alanine substitution mutants between 1522 and 1547 were prepared in a plasmid expressing TET1(1–1547) with the $\Delta 109-120 + \Delta 1132-1202$ ($\Delta 1+2$) mutations (red crosses). (d) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 2$).

NIDs 1 and 2 for NANOG binding. With both wild-type and double-mutant constructs, the TET1-NANOG interaction was dramatically impaired when the TET1 C-terminus was truncated from residue 1547 to 1521 (Figure 4(b)). Smaller fragments (1–1521, 1–1494, 1–1472 and 1–1379) retained a weak residual interaction with NANOG, which was abolished in double mutants ($\Delta 1+2$). These results mapped a third NANOG-interacting domain within TET1 to residues 1522–1547 (NID 3). Most of these residues are strictly conserved in evolution as they are contained within the cysteine-rich catalytic domain (Figure S4(e)). To identify the residues within this region that bind NANOG, a TET1 construct carrying mutations in NIDs 1 and 2 (TET1

$\Delta 1+2$) was further modified by alanine substitution of serine, positively charged or aromatic residues within residues 1522–1547 (Figure 4(c)). The aromatic \rightarrow A construct (F1523A, F1525A, W1529A, Y1532A, F1533A, F1538A, F1547A), but not other mutants, reduced the NANOG interaction to a similar extent as the truncated negative control (TET1 1–1521 $\Delta 1+2$) (Figure 4(d)). These data indicate that aromatic residues within TET1 1522–1547 play a critical role for interacting with NANOG in ESCs. Finally, a full-length TET1 mutant containing mutations in the three NANOG-interacting domains identified in this study ($\Delta 109-120 + \Delta 1132-1202 + 1522-1547$ aromatic \rightarrow A) was generated and tested (Figure 5(a)). Interestingly, the sequential

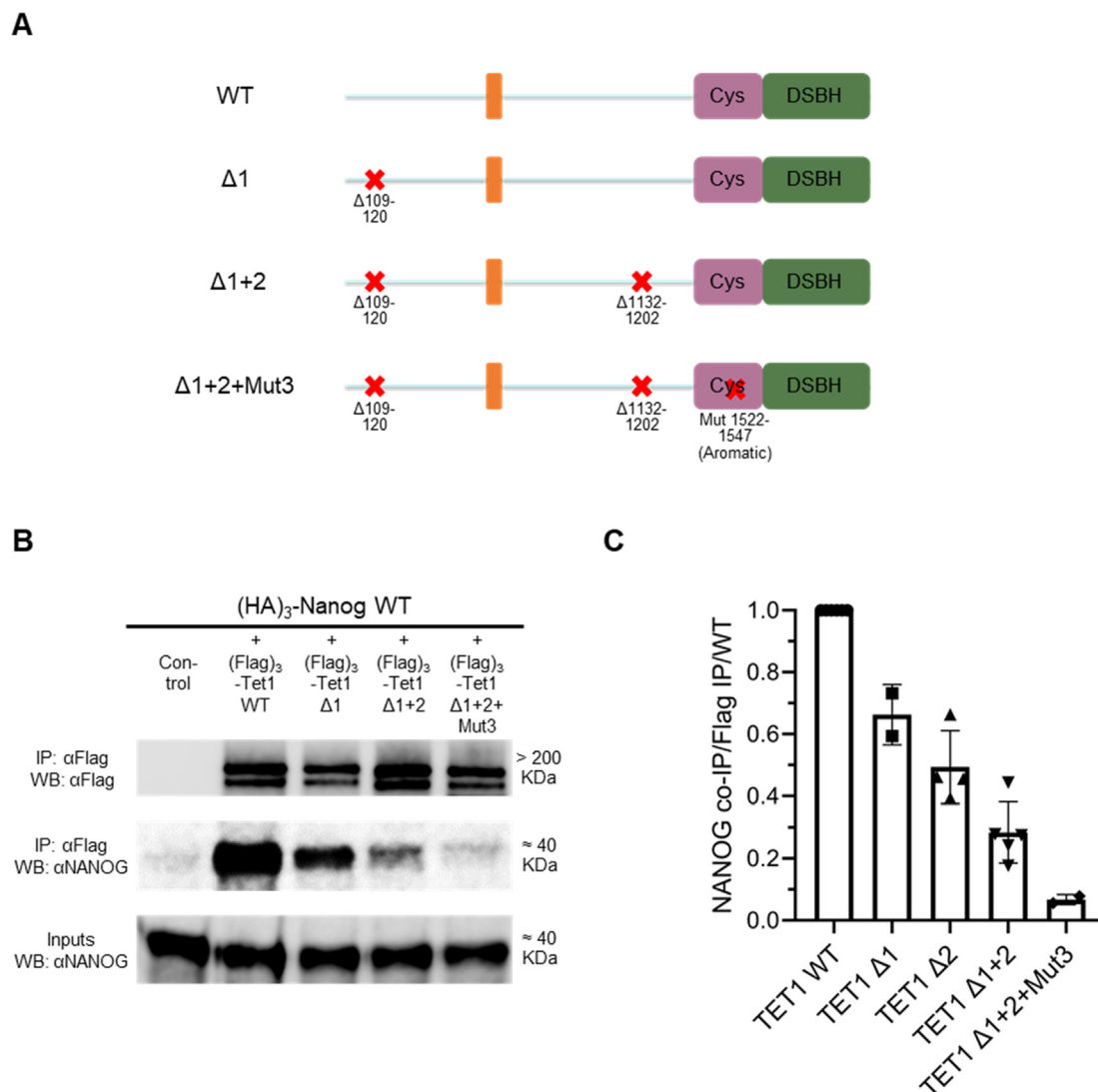


Figure 5. Generation of a TET1 triple-mutant unable to interact with NANOG. Co-immunoprecipitation of the indicated full-length (Flag)₃-TET1 mutants (a) with (HA)₃-NANOG from E14/T ESCs. (b) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 2$). (c) Co-immunoprecipitated NANOG protein was normalised to TET1 immunoprecipitation levels and expressed relative to wild-type; data points indicate independent experiments and error bars standard deviation.

mutations of NIDs 1, 2 and 3 within full-length TET1 gradually decreased the interaction with NANOG, to a level comparable to the negative control (Figure 5(b) and (c)).

To identify the NANOG region(s) interacting with TET1 C-terminus, a series of (HA)₃-tagged NANOG mutants were expressed in ESCs, together with (Flag)₃-TET1 734–2039 (Figure 6(a)). Strikingly, only the NANOG mutant lacking the WR region (NANOG Δ WR) showed a reduced interaction with the TET1 C-terminus (Figure 6(b)). To identify residues within WR responsible for protein–protein interactions, particular amino acids (tryptophan, asparagine, serine and threonine) were substituted by alanine within the WR region of full-length NANOG (Figure 6(c)). Only the W→A mutant showed a decreased interaction with the TET1 C-terminus (Figure 6(d)), demonstrating a key role for tryptophans in the interaction of NANOG with the Tet1 C-terminus. However, full-length TET1 retained its interaction with the W→A mutant (Figure S5), confirming that other NANOG regions interact with TET1 N-terminus. Together, these experiments demonstrate that NANOG interacts with the TET1 C-terminus via aromatic residues conserved in both proteins (Figure 6(e)).

TET1 and NANOG co-bind a subset of pluripotency enhancers associated with NANOG transcriptional target genes

Although TET1 and NANOG interact directly in ESCs, the relationship between the two proteins on chromatin remains unclear. To identify genomic sites potentially regulated by the TET1-NANOG complex, published TET1 and NANOG ChIP-seq datasets were compared. TET1 ChIP-seq peaks from two independent datasets^{23,24} showed an overlap of 13,279 “high confidence” TET1 binding sites (Figure S6(a)). A similar analysis of two NANOG ChIP-seq studies^{38,39} identified 24,357 “high confidence” NANOG ChIP-seq peaks (Figure S6(b)). Subsequently, TET1 and NANOG ChIP-seq signals were visualised at high confidence NANOG and TET1 binding sites, respectively. Interestingly, TET1 is enriched at the centre of a large proportion of NANOG binding sites in ESCs, and this signal is abolished upon Tet1 knock-down (Figure 7(a)). In contrast, NANOG is enriched only at a small proportion of TET1 binding sites in ESCs (Figure S6(c)). Consistent with this low level of co-enrichment, the stringent intersection of “high confidence” TET1 and NANOG ChIP-seq peaks identified only 2003 sites bound by both TET1 and NANOG (Figure S6(d) and (e)). As a first inspection, TET1-NANOG peaks were crossed with relevant genomic features, showing a large proportion of sites corresponding to ESC enhancers⁴⁰ (65%) and a smaller proportion overlapping with CpG islands⁴¹ (22%) (Figure 7(b)). Remarkably, *de novo* motif analysis identified the SOX2/OCT4

composite motif at TET1-NANOG co-bound sites (Figure 7(c)). Following these observations, further analyses were performed to characterise genes associated with TET1-NANOG ChIP-seq peaks. Gene ontology analysis identified groups of genes associated with pluripotency among the top categories, such as “stem cell population maintenance”, “cellular response to leukemia inhibitory factor” and “cell fate specification” (Figure S6(f)). Importantly, TET1-NANOG ChIP-seq peaks were found within or in proximity to 48% of NANOG transcriptional target genes⁴² (Figure 7(d) and Table 1). Visual inspection of these loci showed enrichment of TET1 and NANOG ChIP-seq signals at known enhancers and putative cis-regulatory elements (Figure S6(g)). Together, these results suggest that the TET1-NANOG complex regulates a significant subset of NANOG target genes.

Discussion

TET1¹⁹ and NANOG^{43,44} are both expressed in the inner cell mass of the blastocyst, which is modelled *in vitro* by ESCs. TET1 and NANOG are also co-expressed in the post-implantation epiblast^{20,45,46} and in developing primordial germ cells.^{46,47} Loss of either TET1 or NANOG compromises germline development.^{48–51} Therefore, the TET1-NANOG interaction reported here may function not only at pre-implantation stages but also during later development.

Alanine substitution mutagenesis identified aromatic and hydrophobic residues that mediate the interaction between TET1 and NANOG. Tryptophan residues within the NANOG WR interact with aromatic residues in the TET1 NID 2/3, suggesting an interaction by aromatic stacking. Tryptophans within the WR are also critical for NANOG homodimerization and binding to SOX2 aromatic residues.^{52,53,30,54} TET1 might compete with SOX2 for binding to the WR. Alternatively TET1 and other partner proteins, such as SOX2, could bind simultaneously to different residues within the WR to form larger protein complexes. The present work also demonstrates a direct, WR-independent interaction between NANOG and the TET1 N-terminus, indicative of novel protein interaction sites in NANOG.

TET1 has previously been reported to interact with the SIN3A PAH1 domain by amphipathic helix formation³⁵ and with the O-linked N-acetylglucosamine transferase (OGT) via C-terminal TET1 residues.⁵⁵ However, none of these TET1 residues overlap with the NANOG-interacting domains identified in this study. Other TET1-interacting proteins have been identified.^{29,31,32,36,56–58} These include thymine DNA glycosylase, which binds TET1 through at least two sites.⁵⁶ However, apart from SIN3A and OGT, the residues mediating these interactions have not been defined. The present work demonstrates a

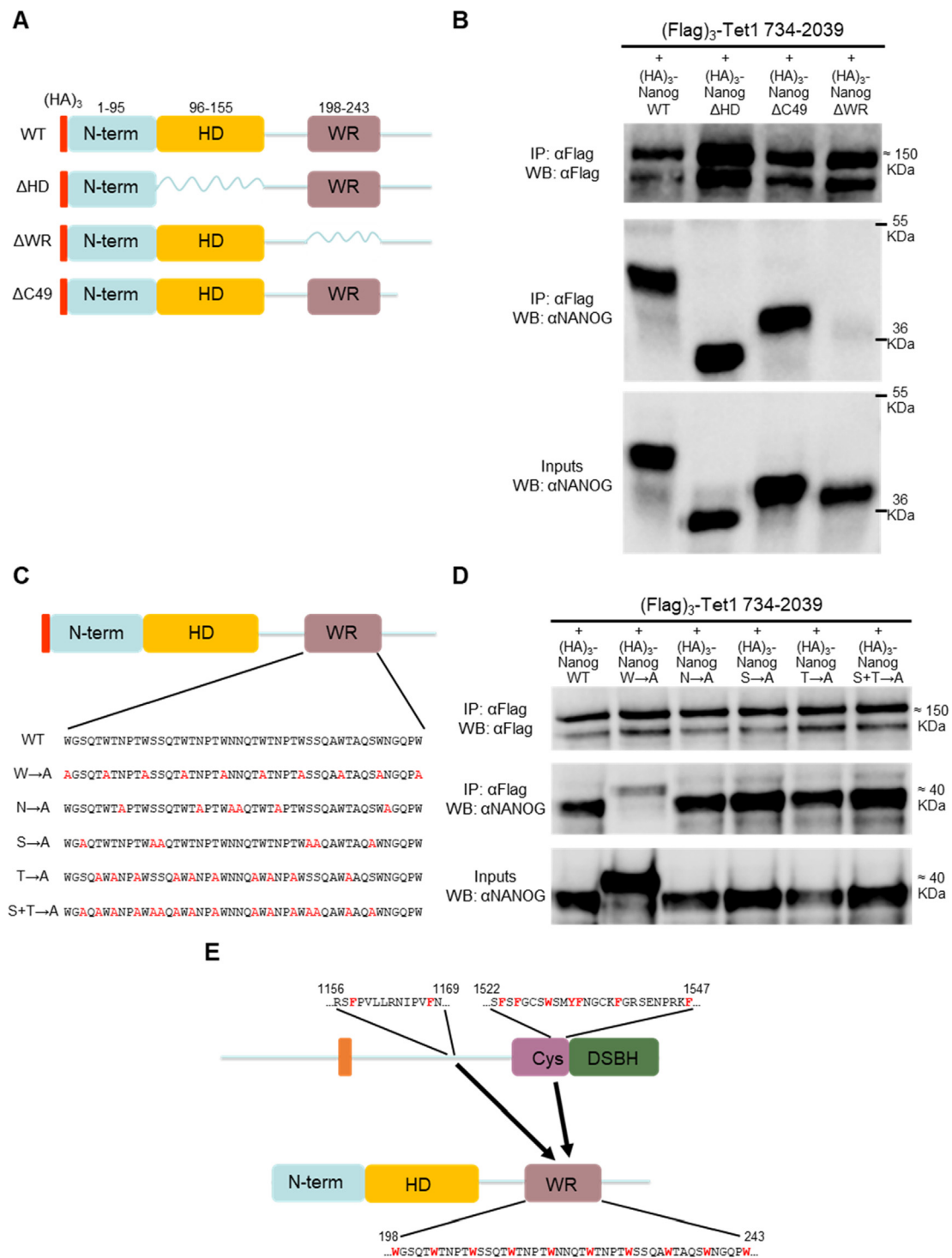


Figure 6. Tryptophan residues within NANOG WR interact with TET1 C-terminus. (a) and (b) Co-immunoprecipitations of (Flag)₃-TET1 (734–2039) with the indicated (HA)₃-NANOG deletion mutants (a) from E14/T ESCs. (b) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 2$). (c) and (d) Co-immunoprecipitation of (Flag)₃-TET1(734–2039) with (HA)₃-NANOG or the derivative alanine substitution mutants within the WR from E14/T ESCs. (c) Only amino acids within NANOG WR region were mutated to alanine. (d) Immunoblots were probed with the antibodies indicated on the left (representative images, $n = 2$). (e) Diagram of the interactions between the TET1 C-terminal domains (NIDs 2 and 3) and NANOG, highlighting aromatic residues critical for the interaction (red).

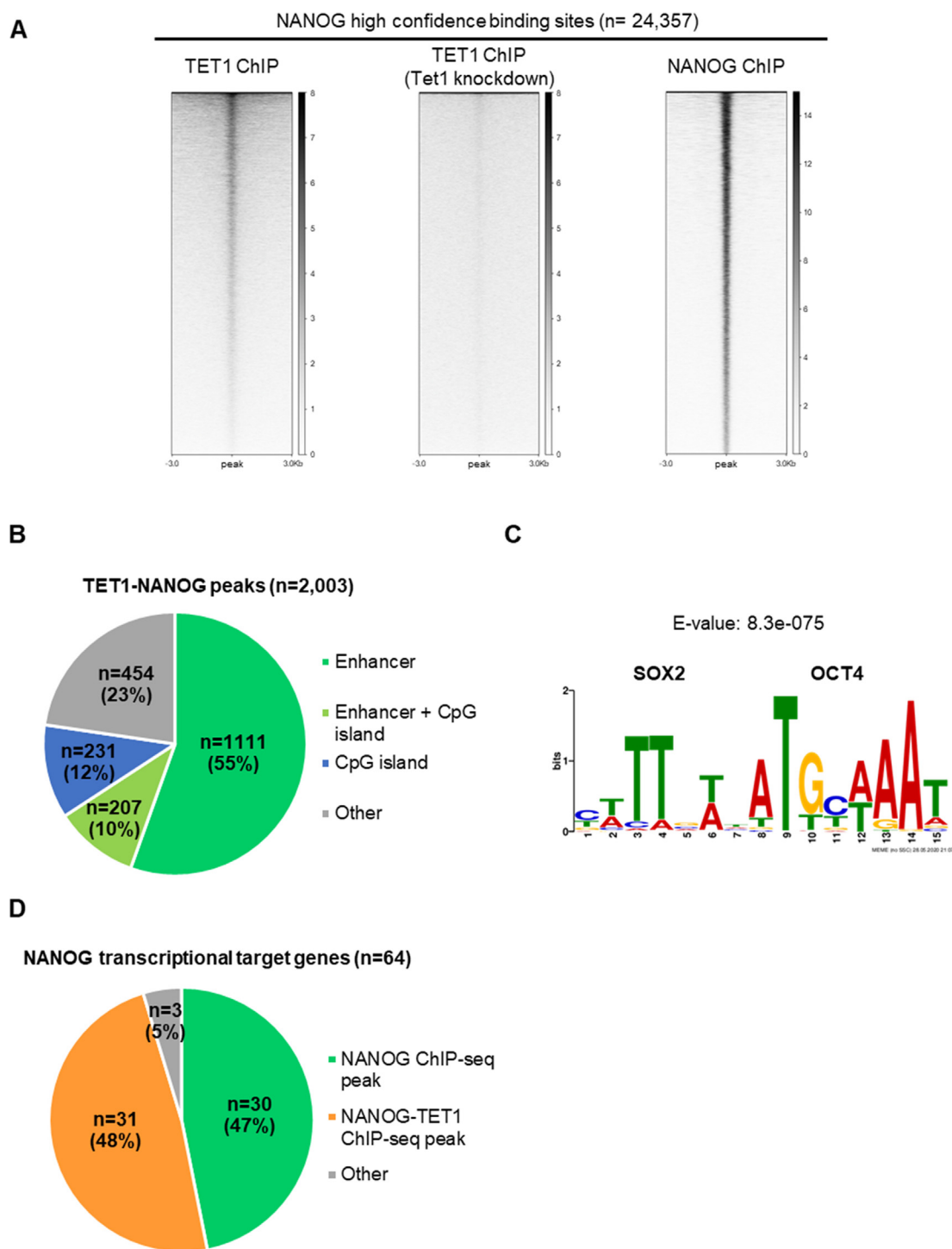


Figure 7. Identification of TET1-NANOG co-bound sites on chromatin in ESCs. (a) TET1 and NANOG ChIP-seq signals at NANOG “high confidence” binding sites, as defined in Figure S6(b). TET1 ChIP-seq in ESCs treated with Tet1 shRNA (knockdown) was used as a negative control. (b) Pie chart showing the portion of TET1-NANOG co-bound sites (see Figure S6(d)) overlapping with ESC enhancers and CpG islands. (c) *De novo* motif analysis performed on TET1-NANOG co-bound sites, showing the most significant binding motif and its respective E-value. (d) Pie chart showing the portion of NANOG transcriptional target genes with NANOG or TET1-NANOG co-bound sites.

direct physical interaction between TET1 and NANOG. Strikingly, the TET1-NANOG interaction involves multiple independent binding regions between both proteins. The three NANOG-

interacting domains on TET1 (residues 109–120, 1156–1169 and 1522–1547) have not been characterised in other protein interaction studies. This adds new information about TET1 function and

Table 1 NANOG transcriptional target genes associated with TET1-NANOG co-bound sites.

Activated genes	Repressed genes
<i>B4gat1</i>	<i>Ctbp2</i>
<i>Cdc42ep4</i>	<i>Dusp1</i>
<i>Cpsf4l</i>	<i>Edn2</i>
<i>En1</i>	<i>Fzd7</i>
<i>Esrrb</i>	<i>Hmces</i>
<i>Igf2bp2</i>	<i>Lefty2</i>
<i>Kit</i>	<i>Lpar6</i>
<i>Klf4</i>	<i>Nid2</i>
<i>Lmo4</i>	<i>Otx2</i>
<i>Manba</i>	<i>Raet1e</i>
<i>Mras</i>	<i>Rbm47</i>
<i>Plekha3</i>	<i>Rbpms</i>
<i>Plpp1</i>	<i>Smagp</i>
<i>Setd1b</i>	<i>Socs3</i>
<i>Sp5</i>	<i>Tcf15</i>
<i>Tmem51</i>	

suggests that TET1 could also interact with other proteins through multiple binding sites. NID 1 (TET1 residues 109–120) binds NANOG in *E. coli*, indicating that the interaction is independent of post-translational modifications. However, the interaction between NID 2 (TET1 residues 1156–1169) and NANOG could not be demonstrated using a bacterial expression system. While this interaction seems to be independent of phosphorylation, other modifications like O-GlcNAcylation^{34,59} might modulate this protein–protein interaction. Interestingly, NID 3 (TET1 residues 1522–1547) includes residues that interact with DNA, that lie adjacent to the TET1 catalytic domain and that are conserved in TET2 and TET3.⁶⁰ We have recently identified two binding regions in TET2 that interact with NANOG and one of these includes residues homologous to NID 3.²¹ Importantly, this region (TET1 residues 1522–1547) also contains residues that bind methylated CpG.⁶⁰ Therefore, binding of NANOG to NID 3 could modulate the interaction of the catalytic domain of TET1 with DNA that depends on these residues. Notably, the TET1-NANOG interaction seems to be DNA-independent since the interaction is seen in DNase-treated protein extracts, and since the interaction is unaffected by deletion of the NANOG homeodomain or the TET1 CXXC domain. It will therefore be of interest to determine whether the interaction with NANOG modulates TET1 catalytic activity.

Here, comparative analysis of TET1 and NANOG ChIP-seq datasets identified a subset of genomic loci co-bound by TET1 and NANOG in ESCs that mainly correspond to pluripotency enhancers. In contrast to most TET1 binding sites that show a broad TET1 ChIP signal, TET1 binding at TET1-NANOG co-bound sites is more narrowly focussed on NANOG peaks. Further supporting NANOG-mediated targeting to these loci, *de novo*

motif analysis of these sites identified the SOX2/OCT4 motif, which characterises NANOG ChIP-seq peaks in ESCs.⁶¹ About half of NANOG target genes have a TET1-NANOG peak nearby, suggesting that TET1 may act cooperatively with NANOG to regulate transcription.⁶² NANOG target genes that have an associated TET1-NANOG peak include genes that are either activated or repressed by NANOG in ESCs. Potentially, TET1 could modulate transcription by demethylating enhancer DNA.^{18,29,63} Alternatively, TET1 may regulate the expression of NANOG target genes by recruiting the SIN3A co-repressor complex at these loci.^{23,64,65} However, further investigation will be required to unravel the mechanisms by which enhancers may be co-regulated by TET1 and NANOG and to distinguish action at positively and negatively regulated NANOG target genes.

Materials and Methods

– Molecular cloning

Mouse TET1 open reading frame was subcloned into pPyCAG plasmids for exogenous expression of (Flag)₃-tagged proteins under a constitutive promoter in embryonic stem cells (see Cell culture section).⁴³ TET1 open reading frame was subcloned into pRSFDuet plasmids (Novagen) for exogenous expression of MBP-tagged proteins under an IPTG-inducible promoter in *E. coli* (see Preparation of protein extracts from bacterial pellets). TET1 truncations and mutants were obtained by cloning PCR products or synthetic DNA fragments (Integrated DNA Technologies, Inc.) using Gibson Assembly.⁶⁶ For more information, please see our list of plasmid constructs, which are available upon request.

– Cell culture

E14/T mouse embryonic stem cells were used in this study, as they constitutively express the polyoma large T antigen and can therefore propagate pPyCAG plasmids carrying the polyoma origin of replication.⁴³ ESCs were cultured in a 37 °C/7% CO₂ incubator on gelatin-coated plates. Composition of the culture medium: Glasgow minimum essential medium (Sigma-Aldrich, cat. G5154), 10% fetal bovine serum, 1× L-glutamine (Thermo Fisher Scientific, cat. 25030024), 1× sodium pyruvate (Thermo Fisher Scientific, cat. 11360039), 1× MEM non-essential amino acids (Thermo Fisher Scientific, cat. 11140035), 0.1 mM 2-Mercaptoethanol (Thermo Fisher Scientific, cat. 31350010), 100U/ml LIF (made in-house).

To overexpress tagged proteins for co-immunoprecipitations, 3 × 10⁶ E14/T ESCs were transfected with pPyCAG plasmids of interest using Lipofectamine 3000 (Thermo Fisher Scientific, cat. L3000008). Transfections were

performed in 10 cm dishes following manufacturer's instructions. E14/T ESCs were harvested 24 h after transfection for protein extract preparation.

– Preparation of nuclear protein extracts from embryonic stem cells

ESCs were washed twice with PBS, trypsinised and pelleted (5 min, 400 g, 4 °C) before lysis in swelling buffer (5 mM PIPES pH8.0, 85 mM KCl) freshly supplemented with 1x protease inhibitor cocktail (Roche, cat. 04693116001) and 0.5% NP-40. After 20 min on ice with occasional shaking, nuclei were pelleted (10 min, 500 g, 4 °C) and resuspended in 1 ml of lysis buffer (20 mM HEPES pH7.6, 350 mM KCl, 0.2 mM EDTA, 1.5 mM MgCl₂, 20% glycerol) freshly supplemented with 0.2% NP-40, 0.5 mM DTT, and 1x protease inhibitor cocktail. The material was transferred into no-stick microtubes (Alpha Laboratories, cat. LW2410AS) and supplemented with 150 U/ml of Benzonase nuclease (Millipore, cat. 71206). Samples were incubated on a rotating wheel for 30 min at 4 °C and centrifuged (16,000 g, 30 min, 4 °C) to remove any precipitate. Nuclear proteins extracts were stored at –80 °C, or used directly for immunoprecipitation or immunoblot. 30–50 µl of protein extract was used as input material and boiled in Laemmli buffer for 5 min at 95 °C.

– Immunopurification of (Flag)₃-tagged proteins from nuclear protein extracts

To immunoprecipitate TET1, 5 µg of anti-Flag (Sigma-Aldrich, cat. F3165) or anti-TET1 (Millipore, cat. 09–872) antibody was added to protein extracts. For negative controls, 5 µg of normal IgG (Santa Cruz) were added to protein extracts. Samples were incubated overnight at 4 °C on a rotating wheel. 30 µl of beads coupled with ProteinA or ProteinG (GE Healthcare 4 Fast Flow Sepharose), previously blocked with 0.5 mg/ml chicken egg albumin (Sigma-Aldrich), were added to each sample, followed by a 2 h incubation at 4 °C on a rotating wheel. Beads were washed 5 times with lysis buffer (20 mM HEPES pH7.6, 350 mM KCl, 0.2 mM EDTA, 1.5 mM MgCl₂, 20% glycerol) freshly supplemented with 0.5% NP-40 and 0.5 mM DTT. Between each wash, samples were centrifuged at 4 °C for 1 min at 2,000 rpm. After the final wash, beads were resuspended in Laemmli buffer and boiled for 5 min at 95 °C.

As an alternative strategy to immunoprecipitate (Flag)₃-tagged proteins, 30 µl of anti-Flag magnetic beads (Sigma-Aldrich, cat. M8823) was added to each protein extract. To immunoprecipitate endogenous TET1 from *Tet1*-(Flag)₃ ESCs, 150 µl of anti-Flag magnetic beads were added to nuclear protein extracts obtained from ~200 million cells, as described above.

Samples were incubated on a rotating wheel for 2 h at room temperature. Following three washes with PBS using a magnet (Thermo Fisher Scientific, cat. 12321D), magnetic beads were resuspended in Laemmli buffer and boiled for 5 min at 95 °C. Samples were stored at –20 °C or analysed directly by immunoblot.

– Preparation of protein extracts from bacterial pellets

Chemically competent BL21(DE3) *E. coli* (NEB, cat. C25271) were transformed with pRSF bacterial expression plasmids of interest. A single colony was inoculated in LB medium supplemented with appropriate antibiotics and incubated overnight in a 37 °C shaker (225 rpm). The overnight culture was diluted (1/50) in a 50 ml flask containing 50 ml of LB medium supplemented with appropriate antibiotics and incubated in a 37 °C shaker (225 rpm) until the culture reached the exponential phase (~3 h, A₆₀₀: 0.5–0.7). 1 mM IPTG was added to the culture to initiate protein expression, and cells were transferred in an 18 °C shaker (225 rpm) for 6 h. Bacterial pellets were collected by centrifugation (5000 g, 10 min) and stored at –20 °C until protein extraction.

To prepare protein extracts, bacterial pellets were resuspended in 5 ml of cold protein extraction buffer (25 mM Tris-HCl pH 8.0, 200 mM NaCl), and sonicated 3x1 min on ice. Samples were centrifuged (16,000 g, 30 min, 4 °C) to remove insoluble material. Bacterial protein extracts were stored at 4 °C or used directly for protein purification. 30–50 µl of protein extract was used as input material and boiled in Laemmli buffer for 5 min at 95 °C.

– Purification of MBP-tagged proteins from bacterial extracts

To purify MBP-tagged proteins, each bacterial protein extract was loaded into a gravity flow column containing 600 µl of amylose resin. The resin was washed once with cold protein extraction buffer (25 mM Tris-HCl pH 8.0, 200 mM NaCl) and MBP-tagged proteins were eluted in 500 µl of cold protein extraction buffer (25 mM Tris-HCl pH 8.0, 200 mM NaCl) supplemented with 10 mM Maltose. 50 µl of eluate was boiled in Laemmli buffer for 5 min at 95 °C.

– Immunoblot

Protein samples were loaded into Bolt 10% Bis-Tris Plus Gels (Thermo Fisher Scientific, cat. NW00102BOX) with 1x Bolt MOPS SDS running buffer (Thermo Fisher Scientific, cat. B0001). 10 µl of SeeBlue Plus2 pre-stained protein

standard (Thermo Fisher Scientific, cat. LC5925) was used to visualize protein molecular weight. The electrophoresis was performed at 160 V for 1 h. Proteins were transferred overnight at 4 °C onto a nitrocellulose membrane (150 mA constant current) with transfer buffer (25 mM Tris, 0.21 M glycine, 10% methanol). The membrane was blocked for 1 h at room temperature with 10% (w/v) non-fat skimmed milk dissolved in PBS supplemented with 0.1% Tween. Then, the membrane was incubated for 1 h at room temperature with primary antibodies diluted to the working concentration in 5% (w/v) non-fat skimmed milk dissolved in PBS supplemented with 0.1% Tween. The membrane was washed 3 times with PBS supplemented with 0.1% Tween, and incubated for 2 h at room temperature with LI-COR IRDye conjugated secondary antibodies diluted 1:5,000 in 5% non-fat skimmed milk dissolved in PBS supplemented with 0.1% Tween. The membrane was finally washed 3 times with PBS supplemented with 0.1% Tween and analysed using the auto-scan function of the LI-COR Odyssey FC imaging system. Molecular weights of protein bands were evaluated by visual comparison with fluorescent protein standards (SeeBlue Plus2 Pre-stained ladder, Thermo Fisher Scientific, cat. LC5925; or Chameleon Duo Pre-stained ladder LI-COR, cat. 928–60000). For protein quantification, the relevant bands were quantified using the LI-COR Image Studio Software.

– Antibodies

Antibody	Reference	Working dilution (immunoblot)
Flag	Sigma-Aldrich, cat. F3165	1:5,000
HA	Covance, cat. MMS-101P	1:5,000
Nanog	Bethyl, cat. A300-397A	1:2,000
GST	Abcam, cat. ab92	1:2,000
His tag	Abcam, cat. ab18184	1:2,000
MBP	NEB, cat. E8032S	1:10,000
Anti-mouse (secondary)	LI-COR, cat. 926-68072	1:5,000
Anti-rabbit (secondary)	LI-COR, cat. 926-32213	1:5,000

– Protein alignments

To identify evolutionary conserved residues, TET1 protein sequences from various mammalian and non-mammalian species were downloaded from UNIPROT (<https://www.uniprot.org/>) and aligned using ESPript (<http://esprict.ibcp.fr>).⁶⁷

– ChIP-seq analysis

ChIP-seq datasets were analysed using the Galaxy platform (<https://usegalaxy.org>).⁶⁸ Details concerning the bioinformatic workflow are available at the following address: <https://usegalaxy.org/u/raf4579/w/workflow-chip-seq-1>. Raw sequencing data (FASTQ files) was downloaded from publicly available databases NCBI's Gene Expression Omnibus or ArrayExpress. Quality control was performed using the software "FastQC" (Babraham Bioinformatics). Samples were filtered to remove contaminating adapter sequences and low-quality reads (cut-off quality score >20.0). Reads were mapped to the mouse mm9 reference genome using "Bowtie2" (BAM file output).⁶⁹ Reads were mapped only to a unique genomic location ($k = 1$). ChIP-seq peaks were called using the software "MACS2" (BED file output).⁷⁰ The immunoprecipitated sample was compared to the genomic input for identifying statistically significant binding sites (qvalue 0.05). For the analysis of NANOG ChIP-seq datasets, the algorithm optimised for "narrow peaks" was used. For the analysis of TET1 ChIP-seq datasets, the algorithm optimised for "broad peaks" was used. If replicates were available, only ChIP-seq peaks shared between replicates were considered for further analyses. Peaks were considered as shared between datasets when presenting an overlap of at least 1 bp. To visualise ChIP-seq datasets on a genome browser, mapped reads (BAM files) from TET1 ChIP-seq with (GSM611195) or without (GSM611194) Tet1 knockdown, as well as NANOG ChIP-seq (GSM1082342) were converted into big-Wig files using "Deeptools".⁷¹ Data was normalised in "Reads Per Kilobase Million" (RPKM) to allow the comparison between ChIP-seq datasets. Genomic snapshots were taken using the genome viewer "IGV".⁷² To visualise ChIP-seq datasets as heat-maps, the software "Deeptools" was used.⁷¹ To perform *de novo* motif analysis on ChIP-seq datasets, the DNA sequences corresponding to each ChIP-seq peak were extracted (FASTA file output) and analysed using the "MEME" software.⁷³ Motifs between 5 and 25 bp, enriched with a *E* value <0.05, were identified. These results were compared to known protein motifs in the JASPAR database.⁷⁴ ChIP-seq peaks were assigned putative target genes using the "Genomic Regions Enrichment of Annotations Tool" (GREAT v4.04: <http://bejerano.stanford.edu/great/public/html/index.php>) with default parameters (basal regulatory domain extending 5 kb upstream and 1 kb downstream from transcription start site, and an extension up to the basal regulatory domain of the nearest upstream and down-stream genes within 1 Mb). Gene ontology analysis from ChIP-seq peaks was performed using the "Genomic Regions Enrichment of Annotations Tool" (GREAT v4.04: <http://bejerano.stanford.edu/great/public/html/index.php>).

Accession numbers

UniProt Knowledgebase (UniProtKB) accession numbers: **E9Q9Y4** (Mouse TET1 protein sequence), **Q80Z64** (Mouse NANOG protein sequence).

NCBI Gene Expression Omnibus accession numbers: **GSE24841** (TET1 ChIP-seq), **GSE26832** (TET1 ChIP-seq), **GSE44286** (NANOG ChIP-seq).

ArrayExpress accession numbers: **E-MTAB-1617** (NANOG ChIP-seq).

CRedit authorship contribution statement

Raphaël Pantier: Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Nicholas Mullin:** Conceptualization, Methodology, Supervision, Writing - review & editing. **Elisa Hall-Ponsele:** Investigation. **Ian Chambers:** Conceptualization, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgements

We thank Elisa Barbieri for constructive comments on the manuscript. We are grateful to Kristian Helin (University of Copenhagen and Memorial Sloan Kettering Cancer Center) for sharing Tet1 expression plasmids. This work was funded by a UK Medical Research Council Grant MR/L018497/1 to IC. RP was supported by a UK Medical Research Council PhD Fellowship.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmb.2020.10.008>.

Received 1 July 2020;
Accepted 7 October 2020;
Available online 13 October 2020

Keywords:

protein–protein interactions;
embryonic stem cells;
chromatin;
pluripotency;
enhancers

Abbreviations:

DSBH, double stranded beta helix domain; ESC, embryonic stem cell; NID, NANOG-interacting domain; OGT, O-linked N-acetylglucosamine transferase; TET, ten-eleven-translocation; WR, tryptophan-repeat

References

1. Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., et al., (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL Partner TET1. *Science*, **324**, 930–935. <https://doi.org/10.1126/science.1170116>.
2. Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300–1303. <https://doi.org/10.1126/science.1210597>.
3. Rasmussen, K.D., Helin, K., (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.*, **30**, 733–750. <https://doi.org/10.1101/gad.276568.115>.
4. Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., et al., (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell*, **13**, 360–369. <https://doi.org/10.1016/j.stem.2013.06.002>.
5. Ficiz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., et al., (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell*, **13**, 351–359. <https://doi.org/10.1016/j.stem.2013.06.004>.
6. Hackett, J.A., Dietmann, S., Murakami, K., Down, T.A., Leitch, H.G., Surani, M.A., (2013). Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Rep.*, **1**, 518–531. <https://doi.org/10.1016/j.stemcr.2013.11.010>.
7. von Meyenn, F., Iurlaro, M., Habibi, E., Liu, N.Q., Salehzadeh-Yazdi, A., Santos, F., Petrini, E., Milagre, I., et al., (2016). Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol. Cell*, **62**, 848–861. <https://doi.org/10.1016/j.molcel.2016.04.025>.
8. Lu, F., Liu, Y., Jiang, L., Yamaguchi, S., Zhang, Y., (2014). Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.*, **28**, 1140–1154. <https://doi.org/10.1101/gad.248005.114>.
9. Hon, G.C., Song, C.-X., Du, T., Jin, F., Selvaraj, S., Lee, A. Y., Yen, C., Ye, Z., et al., (2014). 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell*, **56**, 286–297. <https://doi.org/10.1016/j.molcel.2014.08.026>.
10. Bogdanović, O., Smits, A.H., de la Calle Mustienes, E., Tena, J.J., Ford, E., Williams, R., Senanayake, U., Schultz, M.D., et al., (2016). Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nature Genet.*, **48**, 417–426. <https://doi.org/10.1038/ng.3522>.
11. Rulands, S., Lee, H.J., Clark, S.J., Angermueller, C., Smallwood, S.A., Krueger, F., Mohammed, H., Dean, W., et al., (2018). Genome-scale oscillations in DNA

- methylation during exit from pluripotency. *Cels*, **7**, 63–76. e12. <https://doi.org/10.1016/j.cels.2018.06.012>.
12. Ginno, P.A., Gaidatzis, D., Feldmann, A., Hoerner, L., Imanci, D., Burger, L., Zilbermann, F., Peters, A.H.F.M., et al., (2020). A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. *Nature Commun.*, **11**, 2680. <https://doi.org/10.1038/s41467-020-16354-x>.
 13. Charlton, J., Jung, E.J., Mattei, A.L., Bailly, N., Liao, J., Martin, E.J., Giesselmann, P., Brndl, B., et al., (2020). TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. *Nature Genet.*, 1–9. <https://doi.org/10.1038/s41588-020-0639-9>.
 14. Dawlaty, M.M., Breiling, A., Le, T., Barrasa, M.I., Raddatz, G., Gao, Q., Powell, B.E., Cheng, A.W., et al., (2014). Loss of TET enzymes compromises proper differentiation of embryonic stem cells. *Dev. Cell*, **29**, 102–111. <https://doi.org/10.1016/j.devcel.2014.03.003>.
 15. Verma, N., Pan, H., Doré, L.C., Shukla, A., Li, Q.V., Pelham-Webb, B., Teixeira, V., González, F., et al., (2018). TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nature Genet.*, **50**, 83–95. <https://doi.org/10.1038/s41588-017-0002-y>.
 16. Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., et al., (2013). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell*, **12**, 453–469. <https://doi.org/10.1016/j.stem.2013.02.005>.
 17. Hu, X., Zhang, L., Mao, S.-Q., Li, Z., Chen, J., Zhang, R.-R., Wu, H.-P., Gao, J., et al., (2014). Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell*, **14**, 512–522. <https://doi.org/10.1016/j.stem.2014.01.001>.
 18. Sardina, J.L., Collombet, S., Tian, T.V., Gómez, A., Di Stefano, B., Berenguer, C., Brumbaugh, J., Stadhouders, R., et al., (2018). Transcription factors drive Tet2-mediated enhancer demethylation to reprogram cell fate. *Cell Stem Cell*, **23**, 727–741.e9. <https://doi.org/10.1016/j.stem.2018.08.016>.
 19. Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L. C., Zhang, Y., (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, **466**, 1129–1133. <https://doi.org/10.1038/nature09303>.
 20. Khoueiry, R., Sohni, A., Thienpont, B., Luo, X., Velde, J.V., Bartocetti, M., Boeckx, B., Zwijsen, A., et al., (2017). Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nature Genet.*, **49**, 1061–1072. <https://doi.org/10.1038/ng.3868>.
 21. Pantier, R., Tatar, T., Colby, D., Chambers, I., (2019). Endogenous epitope-tagging of Tet1, Tet2 and Tet3 identifies TET2 as a naïve pluripotency marker. *Life Sci. Alliance*, **2**, e201900516. <https://doi.org/10.26508/lsa.201900516>.
 22. Zhang, H., Zhang, X., Clark, E., Mulcahey, M., Huang, S., Shi, Y.G., (2010). TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res.*, **20**, 1390–1393. <https://doi.org/10.1038/cr.2010.156>.
 23. Williams, K., Christensen, J., Pedersen, M.T., Johansen, J. V., Cloos, P.A.C., Rappilber, J., Helin, K., (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*, **473**, 343–348. <https://doi.org/10.1038/nature10066>.
 24. Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Eve Sun, Y., et al., (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*, **473**, 389–393. <https://doi.org/10.1038/nature09934>.
 25. Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., et al., (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol. Cell*, **42**, 451–464. <https://doi.org/10.1016/j.molcel.2011.04.005>.
 26. Xiong, J., Zhang, Z., Chen, J., Huang, H., Xu, Y., Ding, X., Zheng, Y., Nishinakamura, R., et al., (2016). Cooperative action between SALL4 and TET proteins in stepwise oxidation of 5-methylcytosine. *Mol. Cell*, **64**, 913–925. <https://doi.org/10.1016/j.molcel.2016.10.013>.
 27. de la Rica, L., Deniz, J., Cheng, K.C.L., Todd, C.D., Cruz, C., Houseley, J., Branco, M.R., (2016). TET-dependent regulation of retrotransposable elements in mouse embryonic stem cells. *Genome Biol.*, **17**, 234. <https://doi.org/10.1186/s13059-016-1096-8>.
 28. Ravichandran, M., Lei, R., Tang, Q., Zhao, Y., Lee, J., Ma, L., Chrysanthou, S., Lorton, B.M., et al., (2019). Rinf regulates pluripotency network genes and Tet enzymes in embryonic stem cells. *Cell Rep.*, **28**, 1993–2003.e5. <https://doi.org/10.1016/j.celrep.2019.07.080>.
 29. Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliaha, P.V., Fidalgo, M., Saunders, A., et al., (2013). NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, **495**, 370–374. <https://doi.org/10.1038/nature11925>.
 30. Gagliardi, A., Mullin, N.P., Ying Tan, Z., Colby, D., Kousa, A. I., Halbritter, F., Weiss, J.T., Felker, A., et al., (2013). A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. *EMBO J.*, **32**, 2231–2247. <https://doi.org/10.1038/emboj.2013.161>.
 31. Zhu, G., Li, Y., Zhu, F., Wang, T., Jin, W., Mu, W., Lin, W., Tan, W., et al., (2014). Coordination of engineered factors with TET1/2 promotes early-stage epigenetic modification during somatic cell reprogramming. *Stem Cell Rep.*, **2**, 253–261. <https://doi.org/10.1016/j.stemcr.2014.01.012>.
 32. Okashita, N., Kumaki, Y., Ebi, K., Nishi, M., Okamoto, Y., Nakayama, M., Hashimoto, S., Nakamura, T., et al., (2014). PRDM14 promotes active DNA demethylation through the Ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development*, **141**, 269–280. <https://doi.org/10.1242/dev.099622>.
 33. Zhang, W., Xia, W., Wang, Q., Towers, A.J., Chen, J., Gao, R., Zhang, Y., Yen, C., et al., (2016). Isoform switch of TET1 regulates DNA demethylation and mouse development. *Mol. Cell*, **64**, 1062–1073. <https://doi.org/10.1016/j.molcel.2016.10.030>.
 34. Shi, F.-T., Kim, H., Lu, W., He, Q., Liu, D., Goodell, M.A., Wan, M., Songyang, Z., (2013). Ten-Eleven Translocation 1 (Tet1) is regulated by O-linked N-acetylglucosamine transferase (Ogt) for target gene repression in mouse embryonic stem cells. *J. Biol. Chem.*, **288**, 20776–20784. <https://doi.org/10.1074/jbc.M113.460386>.
 35. Chandru, A., Bate, N., Vuister, G.W., Cowley, S.M., (2018). Sin3A recruits Tet1 to the PAH1 domain via a highly conserved Sin3-Interaction Domain. *Sci. Rep.*, **8**, 14689. <https://doi.org/10.1038/s41598-018-32942-w>.

36. Li, W., Karwacki-Neisius, V., Ma, C., Tan, L., Shi, Y., Wu, F., Shi, Y.G., (2020). Nono deficiency compromises TET1 chromatin association and impedes neuronal differentiation of mouse embryonic stem cells. *Nucleic Acids Res.*, gkaa213. <https://doi.org/10.1093/nar/gkaa213>.
37. Pastor, W.A., Aravind, L., Rao, A., (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature Rev. Mol. Cell Biol.*, **14**, 341–356. <https://doi.org/10.1038/nrm3589>.
38. Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., et al., (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319. <https://doi.org/10.1016/j.cell.2013.03.035>.
39. Karwacki-Neisius, V., G ke, J., Osorno, R., Halbritter, F., Ng, J.H., Weiße, A.Y., Wong, F.C.K., Gagliardi, A., et al., (2013). Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell*, **12**, 531–545. <https://doi.org/10.1016/j.stem.2013.04.023>.
40. Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., Young, R.A., (2013). Super-enhancers in the control of cell identity and disease. *Cell*, **155**, 934–947. <https://doi.org/10.1016/j.cell.2013.09.053>.
41. Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., et al., (2010). Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet.*, **6**, e1001134. <https://doi.org/10.1371/journal.pgen.1001134>.
42. Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., et al., (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell*, **11**, 477–490. <https://doi.org/10.1016/j.stem.2012.08.002>.
43. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., Smith, A., (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*, **113**, 643–655. [https://doi.org/10.1016/S0092-8674\(03\)00392-1](https://doi.org/10.1016/S0092-8674(03)00392-1).
44. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., Yamanaka, S., (2003). The homeoprotein NANOG is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, **113**, 631–642. [https://doi.org/10.1016/S0092-8674\(03\)00393-3](https://doi.org/10.1016/S0092-8674(03)00393-3).
45. Hart, A.H., Hartley, L., Ibrahim, M., Robb, L., (2004). Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev. Dyn.*, **230**, 187–198. <https://doi.org/10.1002/dvdy.20034>.
46. Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N., Tada, T., (2005). Nanog expression in mouse germ cell development. *Gene Expr. Patterns*, **5**, 639–646. <https://doi.org/10.1016/j.modgep.2005.03.001>.
47. Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K., Zhang, Y., (2012). Tet1 controls meiosis by regulating meiotic gene expression. *Nature*, **492**, 443–447. <https://doi.org/10.1038/nature11709>.
48. Yamaguchi, S., Shen, L., Liu, Y., Sendler, D., Zhang, Y., (2013). Role of Tet1 in erasure of genomic imprinting. *Nature*, **504**, 460–464. <https://doi.org/10.1038/nature12805>.
49. Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., et al., (2007). Nanog safeguards pluripotency and mediates germline development. *Nature*, **450**, 1230–1234. <https://doi.org/10.1038/nature06403>.
50. Yamaguchi, S., Kurimoto, K., Yabuta, Y., Sasaki, H., Nakatsuji, N., Saitou, M., Tada, T., (2009). Conditional knockdown of Nanog induces apoptotic cell death in mouse migrating primordial germ cells. *Development*, **136**, 4011–4020. <https://doi.org/10.1242/dev.041160>.
51. Zhang, M., Leitch, H.G., Tang, W.W.C., Festuccia, N., Hall-Ponsole, E., Nichols, J., Surani, M.A., Smith, A., et al., (2018). Esrrb complementation rescues development of Nanog-null germ cells. *Cell Rep.*, **22**, 332–339. <https://doi.org/10.1016/j.celrep.2017.12.060>.
52. Wang, J., Levasseur, D.N., Orkin, S.H., (2008). Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *PNAS*, **105**, 6326–6331. <https://doi.org/10.1073/pnas.0802288105>.
53. Mullin, N.P., Yates, A., Rowe, A.J., Nijmeijer, B., Colby, D., Barlow, P.N., Walkinshaw, M.D., Chambers, I., (2008). The pluripotency rheostat Nanog functions as a dimer. *Biochem. J.*, **411**, 227. <https://doi.org/10.1042/BJ20080134>.
54. Mullin, N.P., Gagliardi, A., Khoa, L.T.P., Colby, D., Hall-Ponsole, E., Rowe, A.J., Chambers, I., (2017). Distinct contributions of tryptophan residues within the dimerization domain to Nanog function. *J. Mol. Biol.*, **429**, 1544–1553. <https://doi.org/10.1016/j.jmb.2016.12.001>.
55. Hrit, J., Goodrich, L., Li, C., Wang, B.-A., Nie, J., Cui, X., Martin, E.A., Simental, E., et al., (2018). OGT binds a conserved C-terminal domain of TET1 to regulate TET1 activity and function in development. *ELife*, **7**, e34870. <https://doi.org/10.7554/eLife.34870>.
56. Weber, A.R., Krawczyk, C., Robertson, A.B., Kuśnierczyk, A., V gbo, C.B., Schuermann, D., Klungland, A., Sch r, P., (2016). Biochemical reconstitution of TET1-TDG-BER-dependent active DNA demethylation reveals a highly coordinated mechanism. *Nature Commun.*, **7**, 10806. <https://doi.org/10.1038/ncomms10806>.
57. Sun, Z., Xu, X., He, J., Murray, A., Sun, M., Wei, X., Wang, X., McCoig, E., et al., (2019). EGR1 recruits TET1 to shape the brain methylome during development and upon neuronal activity. *Nature Commun.*, **10**, 1–12. <https://doi.org/10.1038/s41467-019-11905-3>.
58. Zeng, Y., Yao, B., Shin, J., Lin, L., Kim, N., Song, Q., Liu, S., Su, Y., et al., (2016). Lin28A binds active promoters and recruits Tet1 to regulate gene expression. *Mol. Cell*, **61**, 153–160. <https://doi.org/10.1016/j.molcel.2015.11.020>.
59. Vella, P., Scelfo, A., Jammula, S., Chiacchiera, F., Williams, K., Cuomo, A., Roberto, A., Christensen, J., et al., (2013). Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol. Cell*, **49**, 645–656. <https://doi.org/10.1016/j.molcel.2012.12.019>.
60. Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y.G., Zhu, J., et al., (2013). Crystal structure of TET2-DNA complex: Insight into TET-mediated 5mC oxidation. *Cell*, **155**, 1545–1555. <https://doi.org/10.1016/j.cell.2013.11.020>.
61. Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., et al., (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, **133**, 1106–1117. <https://doi.org/10.1016/j.cell.2008.04.043>.

62. Festuccia, N., Osorno, R., Wilson, V., Chambers, I., (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. *Curr. Opin. Genet. Dev.*, **23**, 504–511. <https://doi.org/10.1016/j.gde.2013.06.003>.
63. Wang, L., Ozark, P.A., Smith, E.R., Zhao, Z., Marshall, S.A., Rendleman, E.J., Piunti, A., Ryan, C., et al., (2018). TET2 coactivates gene expression through demethylation of enhancers. *Sci. Adv.*, **4**, eaau6986. <https://doi.org/10.1126/sciadv.aau6986>.
64. Zhu, F., Zhu, Q., Ye, D., Zhang, Q., Yang, Y., Guo, X., Liu, Z., Jiapaer, Z., et al., (2018). Sin3a–Tet1 interaction activates gene transcription and is required for embryonic stem cell pluripotency. *Nucleic Acids Res.*, **46**, 6026–6040. <https://doi.org/10.1093/nar/gky347>.
65. Saunders, A., Huang, X., Fidalgo, M., Reimer, M.H., Faiola, F., Ding, J., Sánchez-Priego, C., Guallar, D., et al., (2017). The SIN3A/HDAC corepressor complex functionally cooperates with NANOG to promote pluripotency. *Cell Rep.*, **18**, 1713–1726. <https://doi.org/10.1016/j.celrep.2017.01.055>.
66. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., Smith, H.O., (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, **6**, 343–345. <https://doi.org/10.1038/nmeth.1318>.
67. Robert, X., Gouet, P., (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320–W324. <https://doi.org/10.1093/nar/gku316>.
68. Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Eche, M., Chilton, J., Clements, D., et al., (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.*, **44** (2016), W3–W10. <https://doi.org/10.1093/nar/gkw343>.
69. Langmead, B., Salzberg, S.L., (2012). Fast gapped-read alignment with Bowtie 2. *Nature Meth.*, **9**, 357–359. <https://doi.org/10.1038/nmeth.1923>.
70. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D. S., Bernstein, B.E., Nusbaum, C., Myers, R.M., et al., (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137. <https://doi.org/10.1186/gb-2008-9-r137>.
71. Ramírez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dandekar, F., et al., (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160–W165. <https://doi.org/10.1093/nar/gkw257>.
72. Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., (2011). Integrative genomics viewer. *Nature Biotech.*, **29**, 24–26. <https://doi.org/10.1038/nbt.1754>.
73. Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., et al., (2009). tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208. <https://doi.org/10.1093/nar/gkp335>.
74. Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C., Denay, G., Lee, J., Shi, W., Shyr, C., et al., (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.*, **44** (2016), D110–D115. <https://doi.org/10.1093/nar/gkv1176>.